Ambient Noise Classification in the Gulf of Mexico

Mark A. Snyder and Peter A. Orlin Naval Oceanographic Office 1002 Balch Boulevard Stennis Space Center, MS 39529 USA

Abstract - Long-term omni-directional ambient noise was collected in the Gulf of Mexico during 2004 and 2005. The data were recorded continuously and have a bandwidth of 10-1000 Hz. Two-minute averages of Short Time Fourier Transforms (STFT) of the data were computed. The processed data contain wind and wave noise, distant shipping, nearby shipping and storm passage noise with amplitude variation spanning multiple time scales. These contributions to the overall noise level are additive in producing the total measured noise level at any time. An heuristic scheme based on determining the local mean noise level over a period of several hours is used to determine the positive and negative peak values of the data relative to the local mean. The statistics of the negative peaks are used to eliminate large positive peaks (outliers) resulting from nearby shipping. The final local mean is used to characterize long period processes resulting from frontal passages or storms. The removed peak data are used to characterize shipping, and the data with the peaks removed are used to characterize long period ambient noise from wind, waves and precipitation about the local mean. The results of the analysis of the data from January 2005 will be presented.

I. INTRODUCTION

Two principal sources of underwater ambient noise include noise from wind and shipping [1-3]. In deep water, shipping traffic generally dominates in the 10-300 Hz frequency band while weather noise (as measured by wave height or wind speed) generally dominates above 300 Hz. But near shipping lanes, shipping noise can be the loudest source up to 500-1000 Hz. Above 1000 Hz and at higher wind speeds, wind noise usually is the dominant component except for the special cases of nearby or very loud ships [4].

The level of the ambient noise background at a fixed location varies with time. This time variability covers a wide scale, from very fast (short period) phenomena, such as transients of breaking waves, to the very slow (long period) phenomena, such as weather and climate changes [5]. The noise due to shipping varies more rapidly with time than the noise due to the wind. Perrone and King [6] described a technique to classify ambient noise sources as wind-generated or ship-generated by computing the autocorrelation function for different 1/3 octave frequency band time series as recorded by an omni-directional hydrophone. They computed the decorrelation time (the zero-axis crossing time, a measure of the effective width) of the autocorrelation for different frequency bands, and determined that the decorrelation time was 26-40 hours for wind noise but only 4-8 hours for shipping noise.

The distribution of ambient noise values is generally assumed to be Gaussian or near-Gaussian. As Ross [7] pointed out, this offers a good fit to the central distribution, but doesn't accurately represent the tails. The actual distribution appears to be closer to a Rayleigh distribution over certain frequency bands [8,9]. The probability density function (PDF) of the peaks in a zero-mean, narrow-band Gaussian process is a Rayleigh distribution [10].

This information was used to design a noise classification scheme. The goal was to design a system that could count the ships passing within the acoustic range of a hydrophone, and also to be able to determine the passage of storms. Having this information allows an estimate of the ambient noise in an area if all of the local shipping is removed. It also allows an estimation of the ambient noise in an area if significant weather is removed and the estimation of the baseline ambient noise for a region, to which the effects of shipping and extreme weather can be added back in if desired. The classification algorithm is described below.

II. CLASSIFICATION ALGORITHM

Ship passage and weather events are detected and classified using a statistical method based on an assumed distribution of the peaks of the ambient noise in moderately narrowband spectra of the ambient noise. Because ocean ambient noise involves multiple, time varying noise sources which may be probabilistic, deterministic or some combination of the two, the classification method described here is, therefore, heuristic. The method is based on the assumption that ambient noise can be broadly classified into a background level produced by sources which do not leave an identifiable signature in time histories of their spectra ("Background Level"), sources due to wind, rain and wave noise in the vicinity of the sensor ("Weather Level"), shipping ("Shipping Level"), and other sources associated with human produced on-shore and off-shore activities. Each of these sources have their own characteristic frequency and time signatures as well as characteristic occurrences and durations.

maintaining the data needed, and c including suggestions for reducing	completing and reviewing the collection this burden, to Washington Headquald be aware that notwithstanding an	o average 1 hour per response, include ion of information. Send comments rarters Services, Directorate for Information by other provision of law, no person services.	egarding this burden estimate on mation Operations and Reports	or any other aspect of th , 1215 Jefferson Davis I	is collection of information, Highway, Suite 1204, Arlington	
1. REPORT DATE SEP 2007	TE 2. REPORT TYPE			3. DATES COVERED 00-00-2007 to 00-00-2007		
4. TITLE AND SUBTITLE				5a. CONTRACT NUMBER		
Ambient Noise Classification in the Gulf of Mexico				5b. GRANT NUMBER		
				5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S)				5d. PROJECT NUMBER		
				5e. TASK NUMBER		
				5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Oceanographic Office,1002 Balch Boulevard,Stennis Space Center,MS,39529				8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)		
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)		
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited						
13. SUPPLEMENTARY NOTES See also ADM002047. Presented at the MTS/IEEE Oceans 2007 Conference held in Vancouver, Canada on Sep 29-Oct 4, 2007.						
14. ABSTRACT See Report						
15. SUBJECT TERMS						
16. SECURITY CLASSIFIC	17. LIMITATION OF	18. NUMBER	19a. NAME OF			
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified	Same as Report (SAR)	OF PAGES 10	RESPONSIBLE PERSON	

Report Documentation Page

Form Approved OMB No. 0704-0188 For the purpose of classifying these sources using an automated system, it is assumed that ambient noise in excess of the background noise is additive. That is, the ambient noise level is only increased above the background level when weather, shipping or other human produced sources are present. Based on this assumption, the negative going peaks about some mean value are used to characterize the background ambient noise level in the absence of weather events and provide a basis for determining whether positive excursions are of sufficient magnitude to be produced by shipping or other human produced sources.

The basis for classifying ambient noise is total band limited power within 1/3 octave bands derived from data sampled at 2500 Hz, Short Time Fourier Transformed (STFT) to 1 Hz, 1 second bins using a non-overlapping Hann window and boxcar averaged to a 2 minute temporal bin width. The individual 1/3 octave power versus time data are processed to detect events within their bandwidth limits ("Band Event") and later combined to determine whether an event occurred over multiple bands, indicating a "Global Event". The processing steps are summarized in Fig. 1 and detailed in the following paragraphs.

Band event detection is performed in two passes. During the first pass, peaks are identified using a deterministic process based on low probability of occurrences and, initially, long averaging intervals. In the second pass, peaks are detected based on statistical differences between the positive and negative going peaks about a short time local mean.

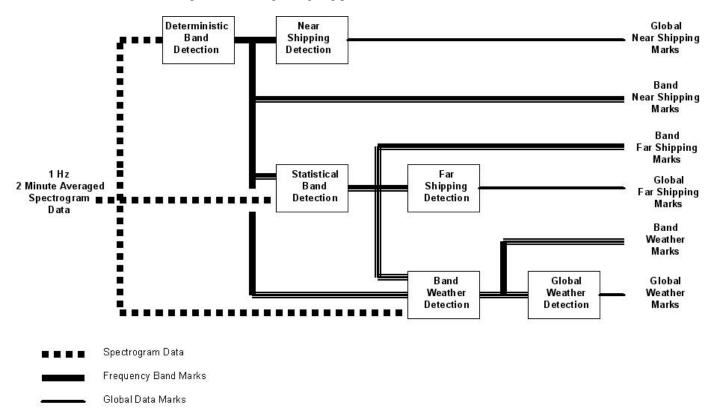


Fig. 1. The processing steps in the classification algorithm.

The detection of a peak, in both cases, is based on the determination of a local mean value of the ambient noise power within a band and the magnitude of a positive peak above the mean relative to the negative peaks about the mean. The mean is calculated using a centered low pass moving average filter evaluated at each 2 minute point in the time series. The positive and negative peaks are found in each interval and the negative peaks are then fit to a Rayleigh distribution under the assumption of a stationary narrowband process within the averaging interval. This procedure is the same for both the deterministic and probabilistic event detection processes. The procedure is performed for each of the bands.

The band-limited power data is processed for peak event detection by detecting low probability positive peaks relative to the negative peaks. The initial averaging period is sequentially decreased from an initial long term value, typically one half a day or more, to an average over 1 hour. Peaks above an unlikely probability, typically starting at 10^{-15} , are marked as near ship passage events in the band. The probability is also decreased, along with the averaging period, from its initial value to a final unlikely

value, typically 10^{-10} . The reference level for the probability is determined from the variance of the negative peaks applied to a Rayleigh distribution.

The deterministic processing is performed independently for each of the bands. At the completion of detection for a given averaging interval and cutoff probability, the peaks exceeding the cutoff probability are eliminated from the data. The data, less the detected peaks, is considered as a single dataset without gaps for subsequent processing. The moving mean is then calculated for the reduced dataset at the current averaging interval. The processing of the band is then repeated at the current cutoff probability level. The processing is completed when no peaks exceeding the probability are found or the ratio of the variance of the positive and negative peaks is minimized. The events detected with deterministic processing are marked as ``near passage'' events for each band.

Each band dataset with the near passage events removed is next processed on a probabilistic basis. The local moving mean value is calculated for a fixed time, typically 1 hour, and the positive and negative peaks are detected. A Rayleigh distribution is assumed for the negative peaks and the probabilities of the positive peaks are divided into an equally spaced set of percentiles, typically ten. The number of peaks within each positive percentile is compared to the number of negative peaks within the same probability range. If the number of positive peaks exceeds the number of negative peaks, the excess positive peaks are then randomly selected for removal. This process is repeated for each probability percentile until the minimum probability value, typically 10⁻¹, is reached. The detected peaks are removed, the mean is recalculated and the peak removal process repeated until the peak counts are balanced or the ratio of the variance of the positive and negative peaks is minimized. The events detected with probabilistic processing are marked as ``far passage" events for each band.

The band events detected by deterministic and probabilistic processing are combined over time within each detection class based on adjacency or near adjacency. Events within a fixed interval, typically 10 minutes, are considered to be part of the same event and all times within the two outer boundaries are marked as an event. This processing compensates for the variability in the power data due to interference and other power "dropout" effects during a ship passage.

The band events for near and far shipping are individually combined into global near and far shipping events if the weighted number of band detections exceeds a specified fraction of the 1/3 octave bands. For near shipping events, this is typically 90% and for far shipping 75%. (Nearby ships have a wider frequency bandwidth received at the hydrophone during the closest point of approach to the buoy than do far ships. This can easily be seen on a spectrogram and is explained by higher frequencies undergoing more propagation loss due to the increased range of the far ships). The weighting for both types of events favors higher frequencies with weights that are proportional to the square of band's width. Near shipping events are typically grouped for global near shipping over the band 10-1000 Hz. Far shipping events are typically grouped for shipping over the band 40-600 Hz.

The global near and far shipping events are individually checked for adjacency or near adjacency. Events within a fixed interval of, typically, 10 minutes, are coalesced in into a single event. Following this processing for the individual global near and far shipping events, adjacency is determined for near and far global events. If a far global event is adjacent or nearly adjacent to a near shipping event, it is declared part of the near shipping event and included with that event. Typically, the near adjacency interval is 10 minutes. This processing is performed to include the edges of near shipping events where the power level is significant but below the criteria for declaring it a near shipping event.

The global events are finally checked for minimum width. If the event is less than the minimum event interval it is excluded from the list of near or far shipping global events. Typically, the minimum width is set to 10 minutes.

Weather events are detected on a band and global basis in a manner similar to ship passage detection. The band power data with near and far shipping events excluded is moving averaged over a relatively long period, typically 48 hours. In contrast to ship passage detection, the power data associated with ship passages is excluded and replaced with power data interpolated between the end points of the ship passages. The filtered data is then least squares fit with a quadratic mean and the mean subtracted from the filtered data. This data is then used to declare whether a weather event has occurred. Typically, a value of 0 dB above the mean is used which appears to correspond, approximately, to a threshold of a 1.5 meters significant wave height or a wind speed threshold of 15 knots. Figures 2 and 3 show the strong correlation between the 48-hour smoothed acoustic power at 250 Hz (after the shipping and the quadratic mean have been subtracted) and the significant wave height data as measured by two National Data Buoy Center (NDBC) moored weather buoys. (Fig. 4 shows the location of the NDBC buoys.) The

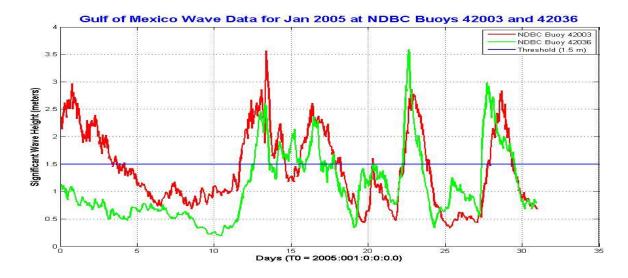


Fig. 2. Wave height data for January 2005 at NDBC buoys 42003 and 42036.

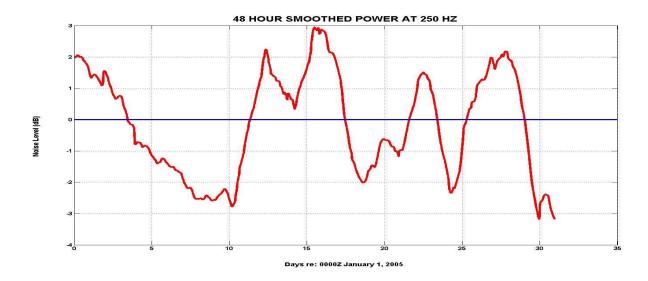


Fig. 3. 48-hour smoothed power for January 2005 at 250 Hz.

detection is qualitative only, since no data for accurate correlation with the magnitude of the power above the detection threshold and wave height at the buoy is available. Unfortunately the two weather buoys were not co-located with the hydrophone.

Following band event detection, the weather events are grouped globally and processed for adjacency following the same scheme as the ship passage events. The 100-1000 Hz band is used for determining a global weather event with a cutoff of 50% of bands indicating that a global event has occurred.

The algorithm produces a set of codes for each data point in each of the frequency bands and a set of codes for each time step for all of the bands combined. These codes identify each point in the time series as part of a global near or far shipping event, a weather event or combined near or far shipping event in weather or data for which no event has been identified. This broadly divides the time history for the spectrogram into undisturbed ambient noise (the background noise level with no ships or significant weather), shipping event levels and weather levels of ambient noise. The coded data can be used by downstream processing to characterize the nature of each noise source.

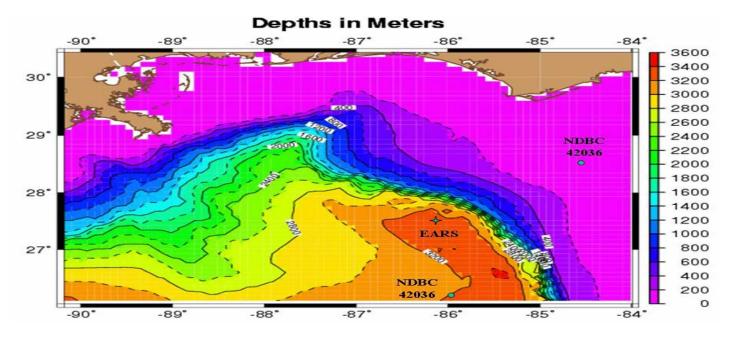


Fig. 4. Gulf of Mexico bathymetry and location of EARS buoy and NDBC weather buoys.

III. RESULTS

The noise classification scheme was used on the acoustic data collected by an Environmental Acoustic Recording System (EARS) buoy in the Gulf of Mexico during January 2005. The bottom-moored EARS buoy was located in a deep portion of the Gulf of Mexico where the water depth was 3200 meters (Fig. 4). The omni-directional hydrophone was located 265 meters above the ocean bottom.

A total of 193 shipping events were detected in 31 days or an average of 6.2 ships per day. The average ship duration was 1.06 hours with a standard deviation of 1.08 hours. The average inter-arrival time (time between ships) was 3.84 hours with a standard deviation of 3.73 hours. Both distributions (ship duration and ship inter-arrival time) were exponential, consistent with a Poisson distribution for shipping. (An exponential distribution has its mean value equal to its standard deviation, which was very close to what was observed for both distributions). When the maximum power of each shipping event was displayed in a histogram, the resulting PDF appeared to be best fit by a Rayleigh distribution.

A total of 4 significant weather events were detected in 31days. The average duration of each weather event was 3.58 days. The average inter-arrival time (time between weather events) was 9.03 days.

As mentioned previously, two weather buoys were near the EARS buoy and provided wind speed and significant wave height data for the same time period. The mean wind speed for the month of January near the EARS buoy was estimated to be 13.0 knots and the mean significant wave height was estimated to be 1.3 meters. These correspond to an average Beaufort Wind Force (BWF) of 4 or an average Douglas sea-state of 4 (SS4). But maximum wind speeds reached about 30.5 knots (BWF 7) and maximum significant wave heights reached about 3.6 meters (SS5) during stormy periods; see Fig. 2. The standard deviation of the wind data was 5.7 knots while the wave height data had a standard deviation of 0.66 meters. The coherence time (time for the autocorrelation to fall to e⁻¹ of its zero-lag value) of the wind time series was 20.0 hours while for the wave height data it was 24.8 hours. (The coherence time turned out to be a better measure of the effective autocorrelation width for these data sets than the decorrelation time described in the introduction to this paper). The wind and wave data had peaks in their power spectra at 56-71 hours (2-3 days), 112-178 hours (5-7 days) and 355-447 hours (15-19 days).

The acoustic data were processed four different ways after using the ambient noise classification scheme. For case 1, all the data were processed. For case 2, shipping traffic was removed. For case 3, significant weather periods (roughly corresponding to sustained wind speeds greater than 15 knots or significant wave heights greater than 1.5 meters) were removed. For case 4, all shipping traffic and significant weather periods were removed. The data removed in each case were based on the global marks. In all four cases, the first four moments of each distribution were computed: the mean, standard deviation, skewness and kurtosis.

The data were examined in eight 1/3 octave frequency bands, centered at the following frequencies: 25, 50, 100, 200, 400, 630, 800 and 950 Hz. In addition, an attempt was made to classify the resulting probability density function (PDF) for each case in each frequency band. The results are shown in Figures 5 to 8.

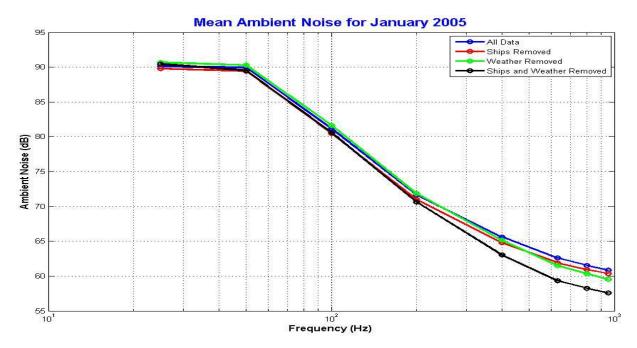


Fig. 5. Mean ambient noise for January 2005.

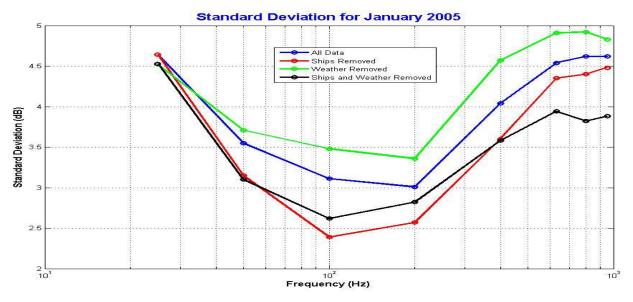


Fig. 6. Standard deviation for January 2005.

Case 1 Results: The complete data set contained 22,320 points (a data point every 2 minutes for 31 days). The mean values (Fig. 5) for the month peaked at 25 Hz and decreased with increasing frequency out to 950 Hz. The standard deviation (Fig. 6) was high at 25 Hz and at higher frequencies (400-950 Hz) but was minimized in the region 50-200 Hz. The skewness (Fig. 7) was positive (skewed towards peaks) from 25-400 Hz but negative (skewed towards troughs) from 630-950 Hz. This was expected. Since shipping noise dominates low frequencies, the region 25-400 Hz was dominated by shipping peaks, which contribute to the high amplitude tails (louder decibel values) of a PDF and make the skewness positive. Weather noise dominates high frequencies, so the region 630-950 Hz was dominated by weather. Weather noise is smoother than shipping noise and

doesn't contribute high level peaks to the data the way nearby ships do. The kurtosis (Fig. 8) was low at 25 Hz, peaked in the region 50-400 Hz and settled down to values near 3 from 630-950 Hz.

At 25 Hz the PDF was bimodal, characterized by two strong peaks in the histogram near 85 and 92 dB. The 25 Hz time series had a strong peak in its fluctuation spectrum at a period of 8 hours. The cause of this 8 hour noise cycle at 25 Hz is being investigated, but it does not appear to be due to shipping or weather.

The PDFs from 50-400 Hz appeared to be best fitted by Rayleigh distributions, to be expected for frequency bands dominated by shipping. These PDFs matched a Rayleigh model quite well for the first and second moments of the distribution (the mean and standard deviation matched, as well as the median and the mode) but not so well for the third and fourth moments (skewness and kurtosis). The PDFs from 630-950 Hz were complicated by the effects of weather and did not fit a simple distribution.

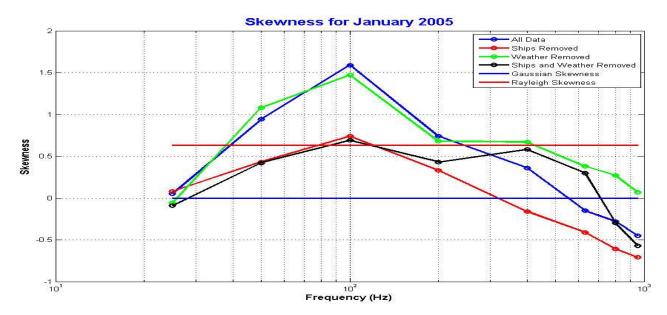


Fig. 7. Skewness for January 2005.

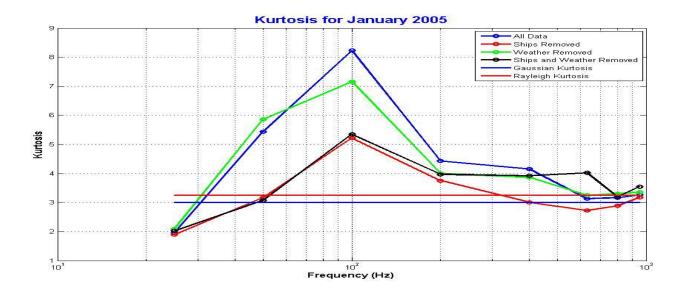


Fig. 8. Kurtosis for January 2005.

Case 2 Results: After shipping traffic was removed, 73.2 % of the original data remained. Since ships show up in time series as short duration spikes, removing them should decrease the mean, standard deviation (less variability), skewness (shrinking the positive tail in the PDF) and kurtosis. This is exactly what was observed. The mean decreased by an average amount of 0.60 dB (all frequency bands showed a decrease). The standard deviation decreased by an average amount of 0.32 dB (all frequency bands showed a decrease except for 25 Hz which was unchanged). The skewness decreased by an average amount of 0.39 (all frequency bands showed a decrease except for 25 Hz which showed a slight increase). The kurtosis decreased by an average amount of 0.99 (all frequency bands showed a decrease, with most of the decrease in the 50 and 100 Hz bands).

All the PDFs for case 2 matched the PDFs for case 1 except at 400 Hz. The 400 Hz PDF no longer appeared to be a Rayleigh distribution because most of the positive tail was removed when the ships were removed.

Case 3 Results: After removing significant weather but not shipping traffic, 62.5 % of the original data remained. Keeping the shipping traffic meant that the positive tails in the PDFs were retained. As compared to case 1, the mean actually increased from 25-200 Hz by an average amount of 0.41 dB but decreased from 400-950 Hz by an average amount of 1.02 dB. Shipping noise is dominant at low frequencies, where the weather contribution to total ambient noise is low and appears to contribute more to the "quiet side" (i.e., the below average side) of the PDF. But at high frequencies, weather noise is dominant, and appears to contribute more to the "loud side" (i.e., the above average side) of the PDF. Removing the significant weather evidently resulted in low dB values being removed at low frequencies but high dB values being removed at high frequencies. Leaving the short duration ship spikes in the data but removing the longer period, less variable weather caused the standard deviation to increase by an average amount of 0.27 dB (all frequency bands showed an increase except for 25 Hz which showed a decrease). The skewness increased by an average amount of 0.22 (3 frequency bands showed a decrease while 5 bands showed an increase, mostly the upper frequency bands). The kurtosis decreased by an average amount of 0.11 (3 frequency bands showed a decrease while 5 bands showed an increase,

As compared to case 1, the 25 Hz band remained a bimodal PDF. The 50, 100, 200 and 400 Hz PDFs remained Rayleigh distributions. But increasing the skewness from 630-950 Hz caused the 630 and 800 Hz distributions to appear to become Rayleigh distributions and the 950 Hz distribution appeared to become Gaussian.

Case 4 Results: After shipping traffic and significant weather were removed, 35.8 % of the original data remained. The mean decreased by an average amount of 1.76 dB (all frequency bands showed a decrease except for 25 Hz which showed an increase). Most of the decrease was in the 400-950 Hz bands. The standard deviation decreased by an average amount of 0.48 dB (all frequency bands showed a decrease). The skewness decreased by an average amount of 0.17 (all frequency bands showed a decrease except for 400 and 630 Hz which showed an increase). The kurtosis decreased by an average amount of 0.58 (four frequency bands showed a decrease while the other four showed an increase). Most of the decrease in kurtosis was in the 50 and 100 Hz bands.

As compared to case 1, the 25 Hz band remained a bimodal PDF. The 50, 100, 200 and 400 Hz PDFs remained Rayleigh distributions. But increasing the skewness at 630 Hz caused the 630 Hz data to appear to fit a Rayleigh distribution. The PDFs for 800 and 950 Hz again did not fit a simple distribution.

A very useful way of displaying the results of all four cases is shown in Figures 9 and 10, which show the exceedance probability (EP) at 16 Hz and 800 Hz, respectively. (The EP for a given frequency band and a specific case 1-4 is related to the corresponding PDF as follows: the integral of the PDF is the cumulative distribution function (CDF) and EP = 1 – CDF.) For example, at 16 Hz, the ambient noise exceeds 80 dB 100% of the time for all four cases. But the median value (50 %) for case 1 (all data) is about 92 dB, meaning the ambient noise exceeds 92 dB 50% of the time. The median value for case 4 (ships and weather removed) is about 86 dB. The case 4 curve on each plot represents the ambient noise baseline (minimum value) for January 2005 for this region in the Gulf of Mexico. Note how removing ships has a huge impact at 16 Hz (compare the blue "All Data" curve with the red "Ships Removed" curve) but very little impact at 800 Hz. Conversely, removing extreme weather has a slight impact at 16 Hz but a much larger impact at 800 Hz. These curves serve as a useful tool in making ambient noise predictions for an area.

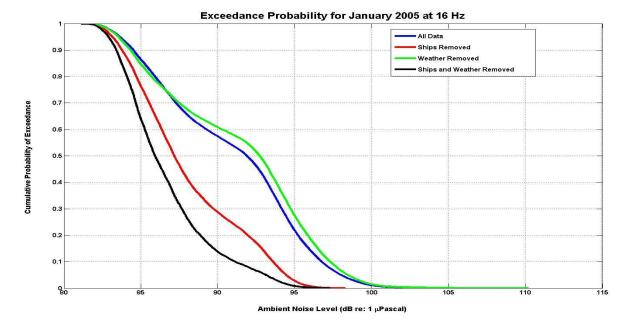


Fig. 9. Exceedance probability for January 2005 at 16 Hz.

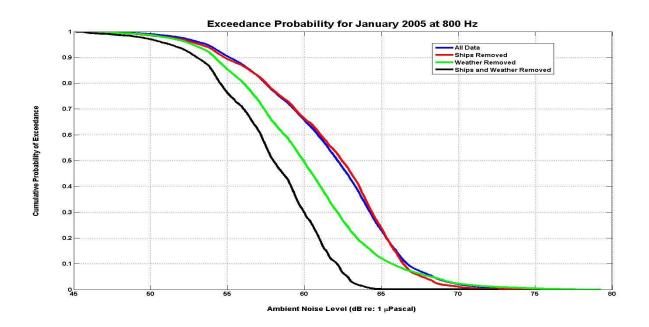


Fig. 10. Exceedance probability for January 2005 at 800 Hz.

IV. SUMMARY

A noise classification method appears useful in estimating ship traffic statistics as well as extreme weather events. The elimination of peaks in the data due to shipping does not appear to significantly affect the gross percentile statistics used to characterize overall ambient noise over long time periods. The elimination of extreme weather events has a larger impact on the gross percentile statistics, especially at higher frequencies. This information allows an estimate of the ambient noise in an area if all of the local shipping and extreme weather periods are removed. It allows the estimation of a baseline ambient noise prediction for a region, to which the effects of shipping and extreme weather can be added if desired. The classification software is currently under development to improve the classification accuracy. Future work will also include a quantitative definition of near and far shipping and an attempt to calibrate the correlation between ambient noise and surface weather.

REFERENCES

- [1] Wenz, G. M., "Acoustic ambient noise in the ocean: spectra and sources," J. Acoust. Soc. Amer. 34, 1936-1956 (1962).
- [2] Wenz, G. M., "Review of Underwater Acoustics Research: Noise," J. Acoust. Soc. Amer. 51, 1010-1024 (1972).
- [3] Urick, Robert J., Principles of Underwater Sound, 3rd Edition. McGraw-Hill, 1983.
- [4] Hall, T., M. Bradley, and L. Emery, "Software Design Description for the Dynamic Ambient Noise Model (DANM), Version 2.0", PSI Technical Report TRS-313, Planning Systems Incorporated, 2004.
- [5] Urick, Robert J., Ambient Noise in the Sea, Naval Sea Systems Command, 1984.
- [6] Perrone, A. J., and L. A. King, "Analysis technique for classifying wind- and ship-generated noise characteristics," J. Acoust. Soc. Amer. 58, 1186 (1975).
- [7] Ross, Donald, Mechanics of Underwater Noise, Peninsula Publishing, 1987.
- [8] Dyer, Ira, "Statistics of distant shipping noise," J. Acoust. Soc. Amer. 53, 564-570 (1973).
- [9] Urick, Robert J., "Models for the amplitude fluctuations of narrow-band signals and noise in the sea," J. Acoust. Soc. Amer. 62, 878-887 (1977).
- [10] Wirsching, Paul H., T. L. Paez, and K. Ortiz, Random Vibrations: Theory and Practice, Dover Publications, p. 164, 1995.